OMHCHO README FILE

Date of this Document: 1 February 2007

Overview

This document provides a brief description of the OMHCHO data product. OMHCHO contains total column HCHO and ancillary information retrieved from OMI global and spatial zoom mode measurements using a retrieval algorithm that is based on non-linear least-squares fitting originally developed for GOME, and adapted for the OMI instrument. In global mode each file contains a single orbit of data covering a swath of approximately 2,600 km wide from pole to pole (sunlit portions only).

Fitting uncertainties for the HCHO slant columns typically range between 40-100%, with the lower end of this range over HCHO hot-spots. This is roughly comparable to what has been achieved from GOME. Uncertainties in the air mass factor (AMF), used to convert slant to vertical columns, are estimated to be 30%. Hence the total uncertainties of the HCHO vertical columns typically range within 50-105%.

Release History and Release-Specific Information

Software Version	1.0.0
ECS Collection Number	2
Public Release	1 February 2007
Validation Release	25 December 2005
Known Issue List	◇ Across-track striping in the data product (see OMSAO_DeStriping_README). Please consult OMSAO_KnownIssues_README for up-to-date information.

Algorithm Description

The algorithm is based on the direct fitting of radiances and irradiances. In particular, and differing from what is commonly referred to as Differential Optical Absorption Spectroscopy (DOAS) fitting, radiances are not divided by irradiances, no logarithms are taken of the spectra, and no high-pass filtering is applied. The three main stages of the algorithm are (1) Solar wavelength calibration, in which the optimum wavelength registration of the solar irradiance measurements is determined and, unless pre-measured laboratory slit function profiles are used (which is the default), the instrument slit function is determined by fitting an asymmetric Gaussian; (2) Radiance wavelength calibration, which finds the optimum wavelength registration for a representative swath of radiance measurements (usually in the middle of the orbit) and determines a common wavelength grid for auxiliary data bases (molecular reference cross sections, *etc*.); and (3) Fitting of all swath lines in the OMI granule. In each stage, the calibration/fitting is performed individually for the 60 cross-track pixels[1] of an OMI swath line. For improved numerical stability, radiances and irradiances are divided by their respective averages over the fitting window; in other words, they are "normalized" to values ~1.

HCHO fitting is performed in the spectral window 325-357 nm, within the UV-2 channel of the OMI instrument. The model that is fitted to the measurements consists of the solar irradiance, attenuated by contributions from HCHO (the target gas), inelastic (rotational Raman, or *Ring*) scattering, and interferences from from other atmospheric gases, including ozone, NO_2, and BrO; it also contains additive and multiplicative closure polynomials and parameters for spectral shift and squeeze, as well as a sampling correction [Chance *et al*., 2005] that is computed on-line. In the current implementation, BrO loading in the fitting window is assumed known from the results of the OMBRO product output. The fit is mostly unconstrained, with the exception of selected parameters, including the spectral shift, which are constrained in order to prevent problems arising from out-of-bounds values.

1 Alternatively: 30 cross-track pixels in rebinned spatial zoom mode, occurring one day per month.

The results from the spectral fitting are HCHO slant columns, which are converted to vertical columns using a look-up table of air mass factors (AMFs). AMF tables have been pre-computed with a radiative transfer model using climatological HCHO profiles and include the effect of clouds (non-scattering, reflecting cloud top). The AMFs used for the conversion from slant to vertical columns are provided in the data product for all ground pixels. For global mode granules, the HCHO retrieval uses cloud fraction and cloud top pressure from the OMI Raman cloud product, OMCLDRR. By default, the HCHO slant-to-vertical conversion is performed with the AMF derived from the OMCLDRR *CloudFraction* and *CloudPressure* data fields, which is stored in the *AirMassFactor* data field in the OMHCHO product file. An additional HCHO AMF, stored in *AirMassFactorO3*, is provided based on the OMCLDRR cloud data fields *CloudFractionforO3* and *CloudPressureforO3*. In future releases this second AMF may become the default since the *forO3* cloud data fields are optimized for (O_3) gas retrievals and are the recommended set of cloud parameters (see the OMCLDRR README). OMCLDRR cloud variables are unavailable for spatial zoom mode granules that occur one day per month. For those cases cloud information is taken from the OMCLDO2 product that derives cloud fraction and cloud top pressure from the O_2-O_2 collision complex.

The algorithm employs several methods to reduce cross-track striping of the HCHO columns. These include outlier screening in the fitting residuals the use of a composite solar spectrum (both employed during the fitting process), as well as a post-processing smoothing of the fitted columns. Particularly the latter method almost certainly introduces an as yet unquantified bias to the fitted columns that the user of the data should be aware of. The smoothed columns are provided in a separate data field, *ColumnAmountDestriped*. For details on all destriping procedures please consult the separate OMSAO_DeStriping_README file.

More details on algorithm specifics can be found in the OMI Algorithm Theoretical Basis Document Vol. 4, and in Kurosu *et al*. [2004].

Data Quality Assessment

Across-track striping (systematically elevated or reduced column values at the same cross track position along the whole track) of the HCHO columns is a presently outstanding issue. This is not unique to HCHO but affects all OMI data products to a higher or lesser degree. Small absorbers like BrO, HCHO and OClO however, are more strongly affected by striping since the column values are of a similar order of magnitude as the stripes, so that the effect is relatively stronger. Various efforts, both at Level 0-1 and 1-2 data processing, are under way to improve this situation, including the method of outlier identification in the fitting residual as employed in the HCHO fit. A satisfactory solution remains still to be found, and users of the HCHO columns provided here must be aware of this issue.

The HCHO data product provides RMS and one standard deviation (1σ) fitting uncertainties, as derived from the fitting covariance matrix. These uncertainties do not include contributions from uncertainties in the measurements or the reference cross sections. In addition to the uncertainties, a fitting diagnostic flag (*FitConvergenceFlag*) provides information on (non-)convergence of the fitting process. This flag should be consulted for more details on the quality of a particular HCHO column datum. For details see the product specification document OMHCHO.fs or consult the File Specification README.

Preliminary Validation

Several validation activities for the OMI HCHO product are ongoing. These include comparisons with other satellite instruments (GOME and SCIAMACHY) and ground-based measurements, as well as comparisons with chemical transport models (GEOS-Chem). At present, preliminary results from satellite comparisons are available.

Direct comparisons with GOME data products are difficult possible since HCHO retrievals from GOME are no longer reliable due to the advanced degradation of the GOME instrument. SCIAMACHY data on the other hand are strongly influenced by instrument artifacts that render the retrieval of HCHO challenging. However, vertical columns of ~2-3•10^{16} mol/cm^2 retrieved from OMI over regions of enhanced formaldehyde (for example, the Ozarks in the U.S.) are in reasonable agreement with what has been observed from GOME in the past and is currently retrieved by SCIAMACHY. The comparisons have shown OMI columns over such hot-spots to be about 30% lower than GOME and SCIAMACHY. All three sensors retrieve similar background values of ~2-4•10^{15} mol/cm^2.

HCHO Sample Images

A number of sample images of global HCHO monthly averages for selected months in 2005 and 2006 can be found on the OMI HCHO Sample Image Page.

Which Data Should Be Used?

Each SAO data product (BrO, HCHO, OClO) contains the data field *MainDataQualityFlag* that should aid the user in the selection of which data to use and which to avoid. Each ground pixel is assigned a value, the range and classification of which are as follows:

Value	**Classification**	**Rationale**
0	Good	All quality checks passed; data may be used with confidence
1	Suspect	Caution advised because one or more of the following conditions are present: • *FitConvergenceFlag* is < 300 (but > 0): convergence at noise level • Column+1σ uncertainty < 0 • Air Mass Factor unreliable because of cloud fraction ≥ 90%
2	Bad	Avoid using data because one or more of the following conditions are present: • *FitConvergenceFlag* is < 0: abnormal termination, no convergence • Column+2σ uncertainty < 0
-1	Missing	No column values have been computed; entries are missing

Product Description

A 2600 km wide OMI swath contains 60 cross-track pixels, ranging in size from 14x24 km^2 (along x across track) in the center of the swath to about 28x150 km^2 at the edges of the swath (median: 15x33 km^2). The pixels on the swath are not symmetrically aligned on the line perpendicular to the orbital plane. However, the latitude and longitude provided with each pixel represents the location of each pixel on the ground to a fraction of a pixel.

The OMHCHO product is written as HDF-EOS5 swath file. A single OMHCHO file contains information retrieved from each OMI pixel over the sun-lit portion of the orbit (a.k.a. an *OMI granule*). The information provided in these files include: Geodetic longitude and latitude, solar and line-of-sight zenith and azimuth angles, total column HCHO with RMS and 1σ fitting uncertainties, longitude and latitude corner coordinates for each OMI pixel, and a range of ancillary parameters that provide information to assess data quality. Average values over an OMI granule for the HCHO total column, uncertainties, and RMS, as well as the percent values of “good” (converged and columns positive within 2σ fitting uncertainties; this includes the “suspect” category from the table above) and “bad” (failed convergence or truly negative columns) provide general, granule-based information on data quality. For a complete list of data fields and their description, please read the file specifications OMHCHO.fs or see the File Specification README.

OMHCHO data are publicly available from NASA's OMI/Aura Data Products Web Page (GES-DISC). Also, subsets of these data over many ground stations and along Aura validation aircraft flight paths are available through the Aura Validation Data Center (AVDC) website to those investigators who are associated with the various Aura science teams.

For questions and comments related to the OMHCHO dataset please contact Thomas P. Kurosu. Please send a copy of your e-mail to Kelly Chance, who has the overall responsibility for this product.

References

OMI Algorithm Theoretical Basis Document, Volume IV, OMI Trace Gas Algorithms, OMI-ATBD-VOL4, ATBD-OMI-04, Version 2.0, August 2002.

K. Chance, T.P. Kurosu, and C.E. Sioris, Undersampling correction for array-detector based satellite spectrometers, Applied Optics 44(7), 1296-1304 (2005).

T.P. Kurosu, K. Chance, and C.E. Sioris, "Preliminary results for HCHO and BrO from the EOS-Aura Ozone Monitoring Instrument", in *Passive Optical Remote Sensing of the Atmosphere and Clouds IV, Proc. of SPIE Vol. 5652* (2004), doi: 10.1117/12.578606.